



Generation Process: Each token uses SOS as initialization context

	[SOS]	The	cat	sits	on	mat
Query [SOS]						
The						
cat						
sits						
on						
mat						